



# Molecular profiling of Greek native germplasm collection of *Rosa canina* L. for enhanced fruit extract production: a comprehensive approach utilizing neutral, gene, and exon-based markers

Anastasia Boutsika · Ifigeneia Mellidou · Katerina Grigoriadou · Katerina Papapanastasi · Nikos Krigas · Eleni Maloupa · Ioannis Ganopoulos · Aliko Xanthopoulou

Received: 13 February 2024 / Accepted: 20 March 2024  
© The Author(s) 2024

**Abstract** The genus *Rosa* L. is globally distributed and *Rosa canina* L. is a distinguished member of multifaceted interest. Apart from the traditional uses of *R. canina* in folk medicine, food and cosmetic industries, or its ornamental applications, its rose hips are renowned for their functional bioactive components. Thus, identifying the genetic diversity within this species is crucial for any plant breeding project. This study employed three molecular markers namely inter simple sequence repeats; (ISSRs), start codon-targeted (SCoTs), and exon-based amplified polymorphisms; (EBAPs) to conduct the first comprehensive genetic analysis of 12 *R. canina* genotypes. DNA extraction, marker selection, and sequences' amplification were performed following established protocols. The resulting genetic data were analyzed for polymorphism, diversity indices, and population structure using various statistical methods including principal component analysis (PCA), unweighted pair group method with arithmetic mean (UPGMA) clustering, and STRUCTURE analysis. The ISSR analysis revealed a high level of polymorphism (81.82%) and identified two major clusters in the UPGMA dendrogram. SCoT and EBAP markers also exhibited

substantial polymorphism (74.56% and 82.11%, respectively) and formed three distinct clusters. PCA indicated a consistent pattern across markers suggesting reliable genetic grouping. STRUCTURE analysis supported the presence of three genetically uniform subpopulations ( $K=3$ ) within the studied *R. canina* germplasm collection. This study provided a comprehensive genetic characterization of the Greek native *R. canina* genebank collection. The observed genetic diversity and population structure offered valuable insights for future breeding programs targeting specific *R. canina* genetic clusters.

**Keywords** Genetic diversity · Breeding · ISSR · SCoT · EBAP · Dogrose · Rosaceae · Rose hips

## Introduction

The genus *Rosa* L. is one of the largest genera in the *Rosaceae* family, with 149 recognized species widespread in Asia, Middle East, Europe, and North America ([www.theplantlist.org](http://www.theplantlist.org), accessed on 1 January 2024). Among ca. 60 wild rose species of Eurasian distribution (Wisseman and Ritz 2007), *Rosa canina* L. (dogrose) is used as ornamental but also in folk medicine, as food or in cosmetics owing either to its aroma and/or the nutraceutical value of its rose hips (Ercisli 2007; Ayati et al. 2018). The latter contain a broad number of high-value bioactive compounds, including vitamin C, flavonoids,

A. Boutsika · I. Mellidou · K. Grigoriadou · K. Papapanastasi · N. Krigas · E. Maloupa · I. Ganopoulos · A. Xanthopoulou (✉)  
Institute of Plant Breeding and Genetic Resources, Hellenic Agricultural Organization—Demeter (ELGO-DIMITRA), Thermi, 570001 Thessalonikis, Greece  
e-mail: axanthopoulou@elgo.gr

tannins, anthocyanins, catechins, carotenoids, fatty acids, tocotrienols, and minerals (Rein et al. 2004; Ayati et al. 2018; Pehlivan et al. 2018; Maloupa et al. 2021). Due to its high nutraceutical importance coupled with only a few allergic reports or side effects to date, *R. canina* is regarded as an important fruit in the food industry which is widely employed to treat inflammations, cold and cough, skin disorders and chronic pain, or is used to prevent diabetes, hypertension, arthritis, aging, cardiovascular diseases etc. (Ahmad et al. 2016; Kerasioti et al. 2019).

The genus *Rosa* exhibits an intricate taxonomy and perplexed evolutionary history due to its varied reproductive strategies and unique pattern of inheritance (Wisseman and Ritz 2007) and DNA fingerprinting has been extensively used to analyze the variability in the genus *Rosa* (e.g., Gahlaut et al. 2021). *Rosa canina* is generally considered as a pentaploid species ( $2n=5x=35$ , with base number of chromosomes  $x=7$ ), with most morphological characteristics as well as flower volatiles being inherited maternally (Wisseman and Ritz 2007). In this context, most microsatellite loci show maternally inherited alleles, whereas most of the random amplified polymorphism design (RAPD) bands are only passed through seeds to the next generation (Wisseman and Ritz 2007). A draft genome of the wild rose *R. multiflora* Thunb. has been recently sequenced serving as a valuable genetic resource for breeding of cultivated roses (Nakamura et al. 2018). Our current understanding of segmental allopolyploid patterns in *Rosaceae* has significantly been broadened due to the sequence of a double haploid rose line from *R. chinensis* Jacq. combining short- and long- reads, as well as after the generation of a high-density single nucleotide polymorphism (SNP) genetic map from a tetraploid *R. ×polliniana* Spreng. (syn. *R. ×hybrida* Schleich.) population (Raymond et al. 2018). Furthermore, transcriptomic research on diverse germplasm, tissues, and types of roses has enabled to date the development of a gene expression database underlying candidate genes of interest for various members of the genus *Rosa* such as those related to pathogen infections (e.g., Dubois et al. 2012; Liu et al. 2018; Neu et al. 2019).

Greece has a broad number of wild-growing *Rosa* species, with native *R. canina* germplasm being one of the most valuable ones as it is associated with high ornamental, pharmaceutical, cosmetic, and nutritional value (Maloupa et al. 2021). The identification of

high-resolving DNA-based markers is of paramount importance to unlock the potential of neglected and underutilized germplasm across the Mediterranean basin (Maloupa et al. 2021). In this context, the present investigation aimed to present the first study on the genetic diversity and relationships of 12 *R. canina* genotypes, by combining molecular characterization strategies based on Inter Simple Sequence Repeats (ISSR) (Reddy et al. 2002), Start codon-targeted (SCoT) (Collard and Mackill 2009) and Exon-based Amplified Polymorphism (EBAP) (Xiong et al. 2022). In particular, the objectives of this study were to: (i) evaluate the genetic diversity and population structure of the *R. canina* genotypes using ISSR as an arbitrary technique, SCoT as a gene-targeting technique markers and EBAP as an exon-targeting technique; and (ii) compare the level of information provided by ISSR, SCoT and EBAP markers to assess genetic similarities among the investigated germplasm. This study provides baseline data on the population structure and genetic information of *R. canina* wild-growing populations from Greece which could be useful and informative for further breeding programs.

## Materials and methods

### Plant materials studied

In total, nine *R. canina* populations were collected from the wider area of the Balkan Botanic Garden of Kroussia (BBGK) in Northern Greece (41°05'27" N, 23°06'36" E), among which seven were wild-growing at 600–650 m of altitude above sea level, and two *ex-situ*-maintained in the BBGK that were originally sourced from the wild (Fig. 1; Table 1). The expeditions were conducted during fall 2021 when the rosehips were ripe and aimed at the selection of genotypes with specific traits such as vigorous growth and strong fruiting potential. For DNA analysis, leaf samples from nine individuals were collected. After taxonomic identification, each genotype was given a unique IPEN (International Plant Exchange Network) accession number by the Institute of Plant Breeding and Genetic Resources (IPBGR) of the Hellenic Agricultural Organization—Demeter (ELGO-DIMITRA). Additionally, another three wild-sourced *R. canina* population samples with vigorous growth and strong fruiting potential were exploited from



**Fig. 1** Partial view of the *ex-situ*-maintained *Rosa canina* germplasm originating from wild-growing Greek native populations employed in this study

**Table 1** Selected *Rosa canina* Greek native genotypes sampled from various sites of Mt Kroussia and other regions in northern Greece assigned with different IPEN (International Plant Exchange Network) accession numbers

IPEN accession number	Greek prefecture	Area	References
GR-1-BBGK-21,261	Central Macedonia	Mt Kroussia	This study
GR-1-BBGK-21,262	Central Macedonia	Mt Kroussia	This study
GR-1-BBGK-21,263	Central Macedonia	Mt Kroussia	This study
GR-1-BBGK-21,264	Central Macedonia	Mt Kroussia	This study
GR-1-BBGK-21,265	Central Macedonia	Mt Kroussia	This study
GR-1-BBGK-21,266	Central Macedonia	Mt Kroussia	This study
GR-1-BBGK-21,267	Central Macedonia	Mt Kroussia	This study
GR-1-BBGK-19,635	Central Macedonia	Mt Kroussia	This study
GR-1-BBGK-14,191	Central Macedonia	Mt Kroussia	This study
GR-1-BBGK-19,674	Western Macedonia	Ziaka	Maloupa et al. (2021)
GR-1-BBGK-03,2229	Epirus	Ioannina	Maloupa et al. (2021)
GR-1-BBGK-19,193	Epirus	Mt Xirovouni	Maloupa et al. (2021)

previous research projects (Maloupa et al. 2021). All collections were performed using the institution's authorized special permit (Permit 26,895/1527 of 21/4/2021); this permit is issued yearly by the Greek Ministry of Environment and Energy after detailed reporting of the applicant.

#### Genetic characterization using ISSR, SCoT and EBAP molecular markers

DNA was isolated from young leaves using the NucleoSpin® Plant II Kit (Macherey–Nagel, Nordrhein-Westfalen, Germany) according to the manufacturer's instructions. DNA concentration and quality were estimated spectrophotometrically at 260 and 280 nm using an Eppendorf BioPhotometer

(Eppendorf, Hamburg, Germany). The integrity of the DNA was determined using gel electrophoresis on a 0.8% (w/v) agarose gel.

After preliminary screening of 50 ISSR (University of British Columbia—UBC), SCoT primers (Collard and Mackill 2009) and EBAP (Xiong et al. 2022), 25 unambiguously scorable and reproducible markers were selected on percentage of polymorphism (7 ISSRs, 9 SCoTs and 7 EBAPs). Oligonucleotide primers complementary to simple sequence repeats (UBC807, UBC810, UBC811, UBC834, UBC835, UBC840 and UBC841) and others complementary to codon targeted polymorphisms (SCoT1, SCoT13, SCoT14, SCoT15, SCoT30, SCoT33, SCoT34, SCoT61 and SCoT66) and exon targeted polymorphisms (EBAP2, EBAP3, EBAP4, EBAP6,

EBAP8, EBAP13 and EBAP21) were used for PCR amplification. The reaction mixture (total volume 25  $\mu$ L) contained the following reagents: 0.5  $\mu$ L dNTPs (10 mM), 2.5  $\mu$ L 10 $\times$ -buffer, 1  $\mu$ L primer (10 mM), 0.5  $\mu$ L template DNA, 0.1  $\mu$ L Taq polymerase (5 U/ $\mu$ L) and 20.3  $\mu$ L sterile dd H<sub>2</sub>O. PCR amplifications were carried out according to Kadoglidou et al. (2023) using an Eppendorf Mastercycler EP Thermal Cycler Range (Eppendorf, Hamburg, Germany): DNA denaturation was initiated at 95 °C for 5 min followed by 35 cycles at 95 °C of 30 s each, for DNA annealing at 48–56 °C (all at 52 °C excepting UBC811 at 48 °C), for annealing of the primers for 90 s and at 72 °C for 90 s for chain extension. The temperature was held at 72 °C for 5 min after the 35 cycles were completed.

The amplification products of the ISSR, SCoT and EBAP markers were separated by electrophoresis on a 1.5% (w/v) agarose gel and stained with ethidium bromide. The size marker was a 2-log DNA ladder (New England Biolabs, Ipswich, MA, USA). The gels were exposed to UV light in a UVItec Transilluminator (UVItec Limited, Cambridge, UK), and UVIDoc software UVIDocMw version 99.04 (UVItec Limited, Cambridge, UK) was used for analysis.

ISSRs, SCoTs and EBAPs alleles were scored based on whether specific fragments were present (1) or absent (0). The phenotypic data matrices for genetic (ISSR + SCoT + EBAP) information were also generated. All matrices were subsequently analyzed in the same fashion. Nei's coefficient (Lynch and Milligan 1994) was determined to assess the genetic variance within the groups of genotypes, while Nei's formula was used to determine their genetic distance (Nei 1978). Principal Component Analysis (PCA) was used as a graphical representation of a matrix to show how closely related the genotypes were. The calculation of number of alleles (Na), effective number of alleles (Ne), Shannon's Information Index (I), Nei's haploid gene diversity (h), Nei's unbiased haploid gene diversity (uh), as well as Percentage of polymorphism (P) was performed using GenAlex 6.5 (Peakall and Smouse 2012) and Microsoft® Excel 2010/XLSTAT©-Pro software (Version 2013.4.07; Addinsoft Inc., Brooklyn, NY, USA).

The Dice similarity coefficient (Dice 1945) as implemented in the ade4 1.7–15 R package (Dray and Dufour 2007) was used along with the Unweighted Pair Group Method with Arithmetic Mean (UPGMA)

clustering algorithm to perform cluster analysis in R 4.0.2 (R Core Team 2020). The ape 5.4–1 package was used to compute bootstrap support (1000 bootstraps), and the phytools 0.7–47 package was used to display the resulting dendrograms (Revell 2012). The 'admixture' and 'independent allele frequencies' models were used to run STRUCTURE 2.3.4 (Pritchard et al. 2000). With a burn-in of 500,000 iterations and 1,000,000 MCMC repetitions for each run, 20 replicates from K=1 to K=6 were used. K was inferred using Evanno's method (Evanno et al. 2005), which was run in the pophelper 2.3.0 R package (Francis 2017). The software Structure threader (Pina-Martins et al. 2017) was used to parallelize distinct runs of K.

## Results

The utilization of ISSR, SCoT and EBAP markers in the genetic analysis of the germplasm collection of *R. canina* provided significant insights into the genetic diversity of the Greek native populations (wild-growing and *ex-situ*-maintained ones sourced from the wild). Regarding the ISSR markers, a comprehensive analysis of 99 genetic loci was conducted on a sample size of 12 individuals (one from each population), yielding a substantial dataset suitable for evaluating their genetic diversity. The average number of distinct alleles (Na) in the *R. canina* germplasm collection was found to be 1.818, whereas the average number of effective alleles (Ne) was 1.524 (Table 2). The Shannon's Information Index (I) was found to be 0.448, suggesting a modest degree of genetic variation. The diversity (h) and unbiased diversity (uh) were determined to be 0.302 and 0.330, respectively, highlighting the comprehensive genetic abundance throughout the germplasm collection. The latter exhibited significant number of polymorphic loci reaching 81.82%, thus outlining the extant genetic diversity. The binary data exhibited a wide range of band patterns with a total of 99 recorded bands, each with a frequency of 5% or higher. Pointedly, we detected no often occurring bands in 25% or 50% or less of the populations, thus suggesting a lack of widespread prevalence in band frequencies.

Accordingly, a comprehensive analysis using SCoT markers was conducted on 114 loci across 12 population samples consisting of the germplasm

**Table 2** Genetic diversity statistics of *Rosa canina* Greek native germplasm collection based on ISSR, SCoT and EBAP markers

Group	ISSR					SCoT					EBAP					
	N	ne	I	uh	P (%)	N	ne	I	uh	P (%)	N	ne	I	uh	P (%)	
<i>Rosa canina</i>	Mean	12	1.524	0.448	0.33	81.82%	12	1.422	0.381	0.275	74.56%	12	1.537	0.457	0.342	82.11%
	SE		0.036	0.025	0.02			0.033	0.024	0.019			0.037	0.025	0.02	

N: sample size; I: Shannon's Information Index; uh: Nei's unbiased haploid gene diversity; P(%): Percent of polymorphism

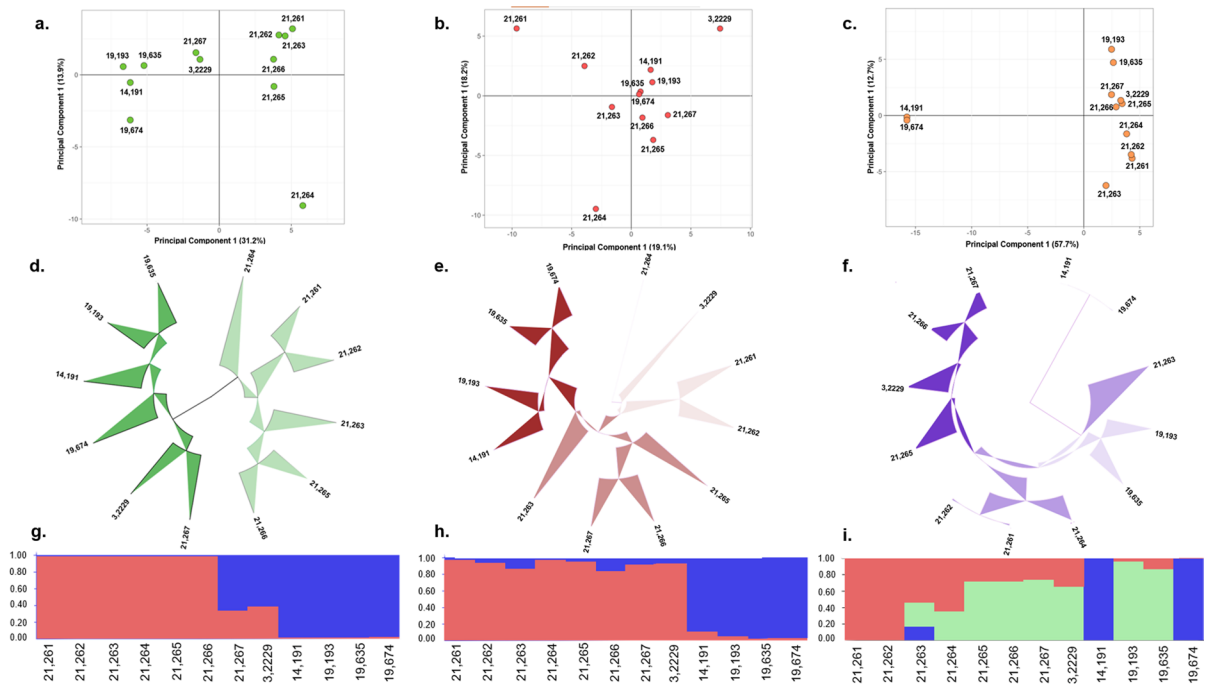
collection. The average number of distinct alleles (Na) was 1.746, whereas the average number of effective alleles (Ne) was 1.422 (Table 2). The Shannon's Information Index (I) was 0.381, suggesting a moderate amount of genetic variation. The diversity (h) and unbiased diversity (uh) were 0.252 and 0.275, respectively. The population exhibited a significant proportion of polymorphic loci amounting to 74.56%, thus emphasizing the presence of genetic diversity within the germplasm collection. The data exhibited significant variation, with a total of 114 recorded bands, all having a frequency of 5% or above. Significantly, no often occurring bands in 25% or 50% or less of the populations were identified, hence suggesting a lack of widespread prevalence in band frequencies.

As for the EBAP markers, a total of 95 loci across 12 population samples were detected. The average number of distinct alleles (Na) was 1.821, whereas the average number of effective alleles (Ne) was 1.537 (Table 2). The Shannon's Information Index (I) was 0.457, suggesting a moderate to high amount of genetic variation. The diversity (h) and unbiased diversity (uh) were 0.309 and 0.342, respectively. The population revealed a significant proportion of polymorphic loci reaching 82.11%, thus highlighting the presence of genetic diversity within the germplasm collection, like other molecular markers. In total, 95 bands were recorded, each with a frequency of 5% or above. Suggestively, no bands that were prevalent in 25% or less or 50% or less of the populations were observed, thus suggesting a lack of general uniformity in band frequencies.

Additionally, UPGMA dendrograms based on Dice distance were constructed for each marker (Fig. 2). The ISSR data revealed the presence of two major clusters contributing to the observed grouping. The initial cluster comprised the genotypes 21,261 to 21,266 (last five digits of IPEN codes), with the remaining accessions being grouped together. The SCoT markers resulted in the formation of three

major clusters; the initial cluster included the genotypes 21,261, 21,262, 21,264, and 3,2229 (last five digits of IPEN codes), the second cluster comprised the genotypes 21,263, 21,265, 21,266, and 21,267, while the last cluster encompassed the genotypes 19,674, 19,635, 14,191, and 19,193 (last five digits of IPEN codes). Concerning EBAP markers, three primary clades were identified once again. The first one comprised of the genotypes 21,261, 21,262, 21,263 and 21,264, the second one consisted of the genotypes 19,635, 19,193, 19,674 and 14,191 genotypes, while the third one included the 21,265, 21,266, 21,267 and 3,2229 genotypes (last five digits of IPEN codes). These findings suggested marker-specific variations in the genetic structure analyses of the studied *R. canina* germplasm collection. The consistent clustering patterns across markers implied reliability in the observed genetic grouping, and the distinct composition of genotypes within each cluster indicated potential genotypic associations and underlying genetic relationships.

The genetic structure was further elucidated by PCA using a covariance matrix and data normalization (Fig. 2). The initial two axes for ISSR, SCoT and EBAP markers accounted for a combined variance of 45.1%, 37.3% and 70.4%, respectively. The samples were evenly distributed over all three markers, showing a consistent pattern without any noticeable segregation. In the context of ISSR markers, a predominant concentration of samples was observed within the upper two quartiles, indicating a potential clustering tendency in the upper range. On the other hand, when considering SCoT and EBAP, a significant grouping of samples was depicted in the rightmost quartiles, indicating a tendency towards higher values in these specific markers. This observation implied a potential correlation between the molecular markers utilized and the distribution pattern of the studied plant material.



**Fig. 2** Cluster analysis and Bayesian inference of genetic differentiation between *ex-situ*-maintained genotypes of wild-growing *Rosa canina* (shown with last five digits of their IPEN accession codes). The results are presented from the left to the right for inter simple sequence repeats (ISSR), start codon-targeted (SCOT) and exon-based amplified polymorphism (EBAP) analysis, respectively. **a. b. c.** Two-D principal com-

ponent analysis (PCA) plot of the first two components for the 12 *R. canina* genotypes; **d. e. f.** Unweighted pair group method with arithmetic mean (UPGMA) dendrograms using Dice distance; **g. h. i.** Genetic assignment based on STRUCTURE using the Values of Evanno's  $\Delta K$  statistic for the most probable genetic structure model

To gain a deeper insight into the genetic composition of the *ex-situ*-maintained *R. canina* germplasm gene pool (a snapshot of the species' Greek native gene pool), STRUCTURE analysis was employed (Fig. 2). The determination of the optimal number of genetic clusters (K) was based on the  $\Delta K$  statistic of Evanno (Evanno et al. 2005), and K=3 was identified as the most fitting value across all molecular markers, including ISSR, SCoT, and EBAP. The latter suggested a consistent and robust genetic structuring pattern within the studied germplasm collection, indicating the representativeness of three distinct genetic subpopulations. The selection of K=3 for all markers highlighted the consensus in determining the genetic composition of the herein studied germplasm and suggested that the *R. canina* collection can be effectively classified into three genetically uniform sections. This systematic method not only improved our understanding of genetic linkages throughout the germplasm collection but also offered useful insights

for future breeding programs and conservation initiatives that focus on specific genetic clusters.

## Discussion

The genetic diversity analysis of the *R. canina* germplasm collection using ISSR, SCoT, and EBAP markers has demonstrated substantial genetic variation within the species. This was consistent with findings from other research using similar markers in *R. canina* and other species. For instance, Jamali et al. (2019) reported that ISSR markers can effectively separate *R. canina* genotypes based on geographic regions, which is crucial for genetic preservation and breeding purposes (Jamali et al. 2019). Similarly, the high genetic diversity observed in our study aligned with findings from Joshi et al. (2021) who has emphasized the potential of ISSR markers in the improvement of the ornamental and perfume industries in

roses (Joshi et al. 2021). The SCoT markers have been shown to be informative in evaluating the genetic diversity and relationships in various species. Previous research has underlined that SCoT markers were reliable in assessing genetic relationships among dill genotypes (Kadoglidou et al. 2023). This was also true for the herein study, where SCoT markers contributed significantly to understanding the genetic makeup of *R. canina* germplasm collection. Xanthopoulou et al. (2015) have also found that SCoT markers were more informative and useful for identifying and analyzing summer squash genetic diversity compared to ISSR markers.

Furthermore, the consistent clustering patterns across the three markers (ISSR, SCoT, and EBAP) indicated robust and reliable genetic grouping within the *R. canina* germplasm collection. These findings have significant implications for sustainable exploitation strategies and future breeding programs. The identification of distinct genetic clusters within the studied germplasm collection represents a valuable resource for breeding programs aimed at enhancing specific traits in *R. canina*. In specific, the genetic makeup of different groups of *R. canina* accessions can facilitate informed decisions about which plants to select in future breeding efforts employing Greek materials; can help identify Greek native biotypes with enhanced levels of bioactive compounds and concomitantly potential health benefits; can detect local biotypes with desirable traits such as high yield potential, enhanced fruit quality, and innate resistance to pests and diseases. Such assets are ultimately aimed to enable future growers to optimize production and meet market demands more effectively.

**Acknowledgements** We thank the personnel of the Institute of Plant Breeding and Genetic Resources of the Hellenic Agricultural Organization Demeter for helping in the cultivation and sampling of the *Rosa canina* germplasm.

**Author contributions** Conceptualization, Aliko Xanthopoulou, Ioannis Ganopoulos; methodology, Anastasia Boutsika, Ifigeneia Mellidou; investigation, Nikos Krigas, Katerina Papanastasi, Eleni Maloupa, Katerina Grigoriadou; data analysis, Anastasia Boutsika; writing—original draft preparation, Aliko Xanthopoulou; writing—review and editing, Anastasia Boutsika, Ifigeneia Mellidou, Nikos Krigas, Katerina Grigoriadou, Ioannis Ganopoulos, Aliko Xanthopoulou. All authors have read and agreed to the published version of the manuscript.

**Funding** Open access funding provided by HEAL-Link Greece. This research was co-funded by the European Union

and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation under the call RESEARCH-CREATE-INNOVATE (project grant number T2EDK-0233; acronym: Green\_Wild\_ROSE.gr).

## Declarations

**Conflict of interest** The authors declare no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ahmad N, Anwar F, Gilani AH (2016) Chapter 76 - Rose hip (*Rosa canina* L.) oils. *Essent Oils Food Preserv, Flavor Saf* <https://doi.org/10.1016/B978-0-12-416641-7.00076-6>
- Ayati Z, Amiri MS, Ramezani M, Delshad E, Sahebkar A, Emami SA (2018) Phytochemistry, traditional uses and pharmacological profile of rose hip: A review. *Curr Pharm Des* 24:4101–4124. <https://doi.org/10.2174/1381612824666181010151849>
- Collard BC, Mackill DJ (2009) Start codon targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. *Plant Mol Biol Rep* 27:86–93. <https://doi.org/10.1007/s11105-008-0060-5>
- Dice LR (1945) Measures of the amount of ecologic association between species. *Ecol* 26:297–302. <https://doi.org/10.2307/1932409>
- Dray S, Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. *J Stat Softw* 22:1–20. <https://doi.org/10.18637/jss.v022.i04>
- Dubois A, Carrere S, Raymond O, Pouvreau B, Cottret L, Rocca A, Onesto JP, Sakr S, Atanassova R, Baudino S et al (2012) Transcriptome database resource and gene expression atlas for the rose. *BMC Genom* 13:638. <https://doi.org/10.1186/1471-2164-13-638>
- Ercisli S (2007) Chemical composition of fruits in some rose (*Rosa* spp.) species. *Food Chem* 104(4):1379–1384. <https://doi.org/10.1016/j.foodchem.2007.01.053>
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol* 14:2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>

- Francis RM (2017) Pophelper: an R package and web app to analyse and visualize population structure. *Mol Ecol Resour* 17:27–32. <https://doi.org/10.1111/1755-0998.12509>
- Gahlaut V, Kumari P, Jaiswal V, Kumar S (2021) Genetics, genomics and breeding in *Rosa* species. *J Hortic Sci Biotechn* 96(5):545–559. <https://doi.org/10.1080/14620316.2021.1894078>
- Kadoglidou K, Cook C, Boutsika A, Sarrou E, Mellidou I, Aidonidou C, Grigoriadis I, Angeli A, Martens S, Georgiadou V, Moysiadis T, Ralli P, Mylonas I, Tourvas N, Michailidis M, Kalivas A, Maloupa E, Ganopoulos I, Xanthopoulou A (2023) Evaluation of a dill (*Anethum graveolens* L.) gene bank germplasm collection using multivariate analysis of morphological traits, molecular genotyping and chemical composition to identify novel genotypes for plant breeding. *PeerJ* 11:e15043. <https://doi.org/10.7717/peerj.15043>
- Liu X, Cao X, Shi S, Zhao N, Li D, Fang P, Chen X, Qi W, Zhang Z (2018) Comparative RNA-Seq analysis reveals a critical role for brassinosteroids in rose (*Rosa hybrida*) petal defense against *Botrytis cinerea* infection. *BMC Genom* 19:62. <https://doi.org/10.1186/s12863-018-0668-x>
- Lynch MMBG, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. *Mol Ecol* 3(2):91–99. <https://doi.org/10.1111/j.1365-294X.1994.tb00109.x>
- Maloupa E, Karapatzak E, Ganopoulos I, Karydas A, Papanastasi K, Kyrkas D, Yfanti P, Nikisianis N, Zahariadis A, Kosma IS, Badeka AV (2021) Molecular authentication, phytochemical evaluation and asexual propagation of wild-growing *Rosa canina* L. (Rosaceae) genotypes of Northern Greece for sustainable exploitation. *Plants* 10(12):2634. <https://doi.org/10.3390/plants10122634>
- Nakamura N, Hirakawa H, Sato S, Otagaki S, Matsumoto S, Tabata S, Tanaka Y (2018) Genome structure of *Rosa multiflora*, a wild ancestor of cultivated roses. *DNA Res* 2:113–121. <https://doi.org/10.1093/dnares/dsx042>
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genet* 89(3):583–590. <https://doi.org/10.1093/genetics/89.3.583>
- Neu E, Domes HS, Menz I, Kaufmann H, Linde M, Debener T (2019) Interaction of roses with a biotrophic and a hemibiotrophic leaf pathogen leads to differences in defense transcriptome activation. *Plant Mol Biol* 99:299–316. <https://doi.org/10.1007/s11103-018-00818-2>
- Peakall R, Smouse PE (2012) GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinform* 28(19):2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>
- Pehlivan M, Mohammed F, Sevindik M, Akgul H (2018) Anti-oxidant and oxidant potential of *Rosa canina*. *Eurasian J Forest Sci* 6:22–25. <https://doi.org/10.31195/ejejfs.475286>
- Pina-Martins F, Silva DN, Fino J, Paulo OS (2017) Structure\_threader: an improved method for automation and parallelization of programs structure, fastStructure and MavericK on multicore CPU systems. *Mol Ecol Resour* 17:268–274. <https://doi.org/10.1111/1755-0998.12702>
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genet* 155:945–959. PMID: 10835412 PMID: PMC1461096
- R Core Team (2020) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing Vienna, Austria. <http://www.r-project.org/index.html>
- Raymond O, Gouzy J, Just J, Badouin H, Verdenaud M, Lemainque A, Vergne P, Moja S, Choisne N, Pont C et al (2018) The *Rosa* genome provides new insights into the domestication of modern roses. *Nat Genet* 50:772–777. <https://doi.org/10.1038/s41588-018-0110-3>
- Reddy M, Sarla N, Siddiq EA (2002) Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. *Euphytica* 128:9–17. <https://doi.org/10.1023/A:1020691618797>
- Rein E, Kharazmi A, Winther K (2004) A herbal remedy, Hyben Vital (stand. powder of a subspecies of *Rosa canina* fruits), reduces pain and improves general well-being in patients with osteoarthritis—a double-blind, placebo-controlled, randomised trial. *Phytomedicine* 11(5):383–391. <https://doi.org/10.1016/j.phymed.2004.01.001>
- Revell LJ (2012) Phytools: an R package for phylogenetic comparative biology (and other things). *MEE* 3:217–223. <https://doi.org/10.1111/j.2041-210X.2011.00169.x>
- Wissemann V, Ritz C (2007) Evolutionary patterns and processes in the genus *Rosa* (Rosaceae) and their implications for host-parasite co-evolution. *Plant Syst Evol* 266:79–89. <https://doi.org/10.1007/s00606-007-0542-1>
- Xiong F, Liu J, Tang R, Yang T, Yang X, He L, Han Z, Qiu L, Zou C, Tang X, Luo C et al (2022) Exon based amplified polymorphism (EBAP): a novel and universal molecular marker for plants. *Electron J Biotechnol* 56:65–74. <https://doi.org/10.1016/j.ejbt.2022.01.001>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.